Operations Research Letters 43 (2015) 132–136

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Topological optimization of reliable networks under dependent failures

Javiera Barrera [a], Héctor Cancela [b], Eduardo Moreno [a,*]

[a] Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Santiago, Chile

[b] Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay

ARTICLE INFO

Article history:
Received 19 May 2014
Received in revised form
28 November 2014
Accepted 25 December 2014
Available online 2 January 2015

Keywords:
Common-cause failure
Dependent failure
Reliable network design
Sample average approximation

ABSTRACT

We address the design problem of a reliable network. Previous work assumes that link failures are independent. We discuss the impact of dropping this assumption. We show that under a common-cause failure model, dependencies between failures can affect the optimal design. We also provide an integer-programming formulation to solve this problem. Furthermore, we discuss how the dependence between the links that participate in the solution and those that do not can be handled. Other dependency models are discussed as well.

1. Problem description

The topological design of reliable telecommunication networks has been deeply studied throughout the last 30 years. The problem can be stated as follows: given a set of nodes and a set of potential links between these nodes, we must choose a subset of links to install such that the total cost is minimized and the reliability is maximized. For a good description of this problem and the different types of reliability requirements, we refer to [12].

The reliability of a network can be defined in different ways. The most common measure of reliability is $\mathcal{K}$-connectivity. That is, given a set $\mathcal{K}$ of terminal nodes, the reliability of the network is the probability that there exists a path from every node in $\mathcal{K}$ to every other node in $\mathcal{K}$. When $\mathcal{K}$ includes all nodes in the network then this measure is called the *all-terminal reliability* and when $\mathcal{K}$ is a specific pair of nodes, it is called the *source-terminal reliability* (or *s-t reliability*). To make the computation of reliability affordable, several simplifications over the failures in a network are made. The most common simplifications are as follows: (a) the link failures are independent; (b) the nodes are perfectly reliable; and (c) no repair is allowed. Note that even under these assumptions, to compute the reliability of a given network is a #P-hard problem [3,20].

Formally, let $\mathcal{G} = (N, E)$ be a graph with node set N and link set E. Let U be a random binary vector taking values in $\{0, 1\}^E$, representing which links are operational. Given U, let E_U be the set of all links that are operational; then, the observed network is given by the graph $\mathcal{G}_U = (N, E_U)$. The $\mathcal{K}$*-reliability* of $\mathcal{G}$ is defined as the probability that $\mathcal{G}_U$ is $\mathcal{K}$-connected.

We study the following problem: given a cost c_e for each $e \in E$ and a budget B, we want to select a subset of links $F \subseteq E$ of total cost less than or equal to the budget, such that the reliability of the selected subnetwork is maximized:

$$\max_{F \subseteq E} \mathbb{P}((N, F_U) \text{ is } \mathcal{K}\text{-connected})$$

$$\sum_{e \in F} c_e \leq B.$$

Due to the difficulty associated with the computation of the reliability of a network, the only known methods to exactly solve this problem are based in the enumeration of the possible solutions [2,9]. This is possible only on small-sized networks and for particular cases, such as when all links have the same probability of failure. Hence, authors have focused on different heuristics and approximation techniques. For example, using Tabu Search [11], Simulated Annealing [19], Genetic algorithms [6,5] or Neural Networks [12,1]. These methods give approximate solutions without any guarantee of convergence to optimality. A recent approach, proposed in [22], employs a sample of failure event scenarios to determine the optimal topology. This technique, called the *sample average approximation* (SAA), converges to an optimal solution when the number of samples is sufficiently large.

* Corresponding author.
E-mail addresses: javiera.barrera@uai.cl (J. Barrera), cancela@fing.edu.uy (H. Cancela), eduardo.moreno@uai.cl (E. Moreno).

http://dx.doi.org/10.1016/j.orl.2014.12.014

However, recent studies question the neglect of the dependence between failures. In [8], authors analyze real data from a Norwegian academic network, showing that neighboring links show significant correlation. In particular, for the design of reliable networks, the impact of dependent failures on the resulting reliability is analyzed in [13], showing that the assumption of failure independence can produce an underestimation of the real reliability. There are several approaches to model the failure dependence between components, including causal failure, cascade failure and common-cause failure. For a discussion of the tractability and scalability of these different models, we refer the reader to [21].

In this paper, we study the topological optimization problem under the best-established common-cause failure model [15]. In Section 2, we present the Marshall–Olkin copula model, which supports common-cause failure dependency. In Section 3, we present an SAA model to solve the problem, and in Section 4, we present some extensions of the previous models to consider other dependent failure models. Finally, we present a computational example in Section 5, showing that ignoring correlation can lead to suboptimal solutions.

2. The Marshall–Olkin copula model for common-cause failures

Common-cause failures are a subset of dependent events in which two or more component fault states exist at the same time and are either direct results or a shared cause [17]. These failures arise naturally in several contexts, for example, in an overlay (virtual) network that is connected through an underlying physical network, so a failure in the physical layer could affect several components of the overlay network. Another example is a network in which the components share equipment that is essential for their function. The Marshall–Olkin (MO) copula introduced in [16] is one of the best-established models for common-cause failures. In this model, events cause one or more components to fail simultaneously, but the lifetime of each link remains exponentially distributed. This model was used in [4] for evaluating the reliability of a network using an importance sampling technique to generate samples of correlated failures.

Formally, let $\mathcal{G} = (N, E)$ be a network. At time zero, we assume that all links are operational. As time passes, links start to fail (alone or simultaneously). Because no repair is allowed, fewer links operate over time. For each link, we can define the lifetime V_e that is the instant at which link e fails. Let $\mathcal{P}_0^E$ be a collection of non-empty subsets of E and let $(W_D)_{D\in\mathcal{P}_0^E}$ be a family of independent positive random variables. The time W_D represents the instant in which a failure that affects all links in D occurs. Therefore, the lifetime of link e is the first time that a set D containing e fails. That is,

$$V_e = \min_{D:e\in D}\{W_D\}.$$

When W_D's are exponential random variables, then this coupling is known as the MO copula. Note that the marginal distribution of the lifetime V_e is also exponential for all e. Let $U_e(t)$ be the state of link e at time t; hence, we have $U_e(t) = \mathbb{I}_{(V_e\leq t)}$. Let $U(t) = (U_e(t))_{e\in E}$; then the graph $\mathcal{G}_U(t) = (V, E_{U(t)})$ is the observed network at time t. Because we are interested in a static model, we take a snapshot of the network at time 1 and evaluate the reliability at that instant (this can also correspond to a dynamic model with a fixed mission time). We fix the status of each link $U_e := U_e(1)$.

Note that this is a natural extension of the independent failure model, setting $\mathcal{P}_0^E$ as the collection of singleton sets, and V_e as exponential random variables of parameter $\log(1-p_e)$. In this particular case, we have independent failure probabilities between links, and at time 1, the failure probability is p_e for link e. We can also recover the model in which the nodes and links fail independently, adding to the previous model the collection of sets $\{D_n : n \in N\}$ such that D_n is the set of all links with n as an end node. In this case, if all nodes fail independently with probability q and the links fail independently with probability p, then the marginal probability failure for each link is $(1-p)(1-q)^2$, and the node failures induce a correlation between adjacent links of $\frac{q(1-p)}{1-(1-p)(1-q)^2}$.

3. An integer programming model using sample average approximations

In this section, we present an integer programming model that considers the dependencies between links, given by an MO copula. This model is related to the ideas of [22], which is based on a sample average approximation (SAA) of the probability that the network is $\mathcal{K}$-connected.

SAA is a popular technique for approximating the stochastic objective function. Some of the first applications of this technique appeared in [10], and the approach was empirically studied in [14]. Recalling that the probability of an event is equal to the expected value of the indicator function of the event, the basic idea is to approximate the expected value by its sampled average. That is, we sample a set of scenarios S, and we approximate the objective function by

$$\mathbb{P}\left((N, F_x) \text{ is } \mathcal{K}\text{-connected}\right) = \mathbb{E}\left(\mathbb{I}_{(N,F) \text{ is } \mathcal{K}\text{-connected}}\right) \approx \frac{1}{|S|}\sum_{s\in S}\mathbb{I}_{(N,F_x^s) \text{ is } \mathcal{K}\text{-connected}},$$

where F_x^s is the subgraph obtained after removing from F_x the failed links in scenario s.

To formulate an integer programming model that uses the previous approximation, we define a binary variable w_s that indicates the event that the graph (N, F_x^s) is $\mathcal{K}$-connected (or not) under scenario s. A $\mathcal{K}$-cut is a subset $M \subseteq N$ such that $\mathcal{K} \cap M \neq \emptyset$ and $\mathcal{K} \setminus M \neq \emptyset$. Recall that a graph $G = (N, E)$ is $\mathcal{K}$-connected if and only if for every $\mathcal{K}$-cut M the cut-set induced by M, $\delta(M) = \{uv \in E : u \in M, v \notin M\}$, is not empty. Finally, let $\{W_D^s : D \in \mathcal{P}_0^E, s \in S\}$ be a sampling of the MO copula. Then, we solve the following integer programming problem:

$$\max \sum_{s\in S} z_s \tag{1}$$

$$\sum_{e\in E} c_e x_e \leq B \tag{2}$$

$$\sum_{e\in\delta(M)} u_e^s \geq z_s \quad \forall M\ \mathcal{K}\text{-cut},\ \forall s \in S \tag{3}$$

$$u_e^s \leq x_e \quad \forall e \in E,\ \forall s \in S \tag{4}$$

$$u_e^s \leq 0 \quad \forall e \in D,\ \forall s \in S \text{ such that } W_D^s < 1 \tag{5}$$

$$x_e \in \{0, 1\} \quad \forall e \in E$$

$$z_s \in \{0, 1\} \quad \forall s \in S$$

$$u_e^s \in \{0, 1\} \quad \forall s \in S,\ \forall e \in E.$$

Binary variables x_e for each link $e \in E$ determine the resulting network, where constraint (2) bounds the total cost of these links. Variables u_e^s represent whether or not link e is operative under scenario s. To do so, constraints (4) and (5) force $u_e^s = 0$ in the case that link e is not chosen ($x_e = 0$), or whether one of the copula associated to link e indicates that this link fails in scenario s ($W_D^s < 1$). Finally, constraint (3) verifies whether the resulting network is $\mathcal{K}$-connected under scenario s or not. To verify this, note that binary variables z_s can take a value of 1 (which is desired by the objective function), only if for every $\mathcal{K}$-cut there is at least one link operative in scenario s. Note that it is possible to eliminate variables u_e^s by replacing constraints (3)–(5) by $\sum_{e\in\delta(M)} \mathbb{I}_{(W_D^s\geq 1)} x_e \geq z_s$, obtaining a similar model to the one presented in [22]. However, we adhere to the model including explicitly the u_e^s variables because we will discuss some extensions in which these variables are required in the model.